REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	3/26/01	Final; 8/1/99-1/31/01

4. TITLE AND SUBTITLE

Elucidating mechanisms of bacterial response

5. FUNDING NUMBERS

DAAD19-99-1-0343

6. AUTHOR(S)

John Cavanagh

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Purdue University, North Carolina State University

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 40293.8-LS

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

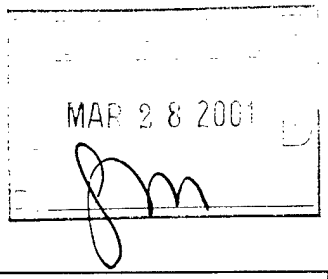
12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for publ ic release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The overall goal of this project was to determine whether protein flexibility contributed to molecular recognition and binding events for proteins critical in the development of bacterial virulence. Such a correlation would provide for an as yet untapped focal point for anti-infective design.



14. SUBJECT TERMS

protein flexibility bacterial virulence
anti-infective design

15. NUMBER OF PAGES

9

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Enclosure 1

20010413 142

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

☐ Reprint (Orig + 2 copies)
☐ Technical Report (Orig + 2 copies)
☐ Manuscript (1 copy)
☒ Final Progress Report (Orig + 2 copies)
☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: DAAD19-99-1-0343

REPORT TITLE: Elucidating Mechanisms of Bacterial Response

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

John Lark

Elucidating Mechanisms of Bacterial Response

Final report for ARO Award DAAD19-99-1-0343

Professor John Cavanagh

Department of Biochemistry

North Carolina State University

Raleigh, NC 27695

919-513-4349

john_cavanagh@ncsu.edu

This research program was terminated by ARO after 17 months rather that running its three year expected course for reasons that are unclear.

Introduction

The overall goal of this project was to determine whether protein flexibility contributed to molecular recognition and binding events for proteins critical in the development of bacterial virulence. Such a correlation would provide for an as yet untapped focal point for anti-infective design.

The crucial class of proteins that form the core of our investigations are called **transition-state regulators.** When bacteria encounter stress, an inevitable event on their daily life course, they respond and adapt in a clear effort to survive. However they do not immediately leap into the purely protective phase as this requires an enormous commitment in energy. To do this entails a complete reworking of the cellular morphology and overall lifestyle. It is not a process the cell enters into frivolously. Rather, the bacteria move into a transition state. This state can be regarded as a cellular holding pattern where process required for both growth and survival are regulated simultaneously. A myriad of processes are controlled during this period. Instead of a single control element for each pathway bacteria have evolved to have regulatory 'generals' during this period. These are proteins that are responsible for the activation or repression of a variety of transcriptional mechanisms. These **transition-state regulator** proteins are able to perform their array of tasks by being able to target and bind to a host of gene promoters which share no consensus oligonucleotide base-pairing sequence.

What mechanism makes them so promiscuous?

To answer this question we chose as our model system the transition-state regulator protein **AbrB** *from Bacillus subtilis* which has been well characterized genetically.

Below we highlight the structural and dynamic investigations of the DNA-binding domain of AbrB and the generation of a model for its promiscuous DNA-binding properties.

Structural Studies of the Transition State Regulator AbrB

Abstract: **The high resolution NMR solution structure of the novel DNA-binding domain of the *Bacillus subtilis* transition-state regulator AbrB is described. Comparisons of the AbrB DNA-binding domain with DNA-binding proteins of known structure show that it is a member of a completely novel class of DNA-recognition folds employing a dimeric topology for cellular function. This new DNA-binding conformation is referred to as the *looped-hinge helix fold*. Sequence homology investigations show that this DNA-binding topology is found in other disparately related microbes. Combination of a structural analysis of the AbrB DNA-binding domain with bioanalytical and mutagenic data of full length AbrB allows for a general model that describes genetic regulation properties of AbrB.**

Transition-state regulators recently have been recognized as a discreet category of proteins that exert regulatory control of protein expression during post-exponential growth. Although much effort has focused on the genetic characterization of transition-state regulators, nothing is known about their structure and function at the molecular level. Understanding the molecular mechanisms of how these proteins function is vital since they allow bacteria, including pathogenic species, to cope with and adapt to changing environments. In *Bacillus subtilis*, transition-state regulators play an essential role in the adaptive capacity and survival of the cell. When environmental conditions are hazardous or non-conducive for exponential growth, alternative metabolic pathways are established to optimize the use of available resources or to direct the cell towards formation of a protective dormant spore. The *B. subtilis* transition-state regulator AbrB is involved in the control of over 60 different genes nominally expressed or repressed in sub-optimal environmental conditions characterized by limited nutrient resources.

AbrB is a multimeric protein consisting of identical 94 residue, 10.5 kDa monomers. Mutational studies have shown that each monomer can be divided into two functional domains. The role of the C-terminal domain (residues 54-94) is in determining the state of AbrB macromolecular assembly. Mutational disruption of the C-terminus alters, but does not eliminate, DNA binding affinity, and results in AbrB mutant solution sizes half that of wild type AbrB. The N-terminal domain (residues 1-53) is primarily involved in DNA binding. Mutations at the AbrB Arg 23 and 24 positions result in no detectable DNA binding activity to a natural DNA promoter sequence, and at the Gly 27 position shows decreased binding activity.

DNase I footprinting analysis on AbrB-bound promoter elements has failed to reveal an obvious DNA consensus sequence recognized by AbrB. However, previous studies have shown that AbrB binding is specific, even in the presence of excess non-specific competitor DNA, and generally occurs in the region immediately upstream or downstream from the target gene +1 initiation site. It also has been demonstrated that regulatory disadvantages are conferred by mutagenesis of natural promoter sites that approximate an optimized, high-affinity *in vitro* selected DNA sequence. This implies that DNA targeting by AbrB depends, in part, on recognition of a general DNA tertiary structure. Elucidation of AbrB structure-function relationships pose a complex challenge in that a general mechanism allowing for promiscuity in DNA target recognition may be inherent primarily to the unbound protein or result from specific protein/DNA interactions unique to each target promoter sequence. How AbrB is able to specifically bind many unrelated genes is previously unknown and prompted these investigations into the structural basis of DNA recognition.

In this work, we elucidated the high-resolution solution structure of the dimeric AbrB N-terminal DNA-binding domain (AbrBN53; residues 1-53), and developed the *in vitro* analysis of AbrB and AbrBN53 solution states and DNA binding capacity. Our results show that nucleotide sequence recognition relies on a novel, dimeric topology of the N-terminal DNA binding domain. Nucleotide sequence discrimination may be a function of both the tetrameric and dimeric state of full length AbrB macromolecular

assembly. These data support a model in which AbrB binds target DNA sequences using an integrated mechanism of macromolecular assembly and reorientation of key features of the novel DNA binding domain.

Dynamics Studies of the Transition State Regulator AbrB

Abstract: **AbrB is a *Bacillus subtilis* protein responsible for regulating a diverse array of unrelated genes during periods of sub-optimal growth conditions. DNA binding by AbrB is unique in that sequence recognition is specific, yet no obvious consensus sequence of bound promoter regions is apparent. The N-terminal domain is a recently characterized representative of a novel class of DNA-binding proteins that possess a looped-hinge helix DNA-binding topology (see above). Although the structural characterization of this DNA-binding topology contributed to an understanding of the architectural basis for recognition of DNA target sequences, specific mechanisms responsible for promiscuity in DNA sequence recognition still were not apparent. Analysis of ^{15}N backbone relaxation parameters shows that dynamic motion of regions directly linked to DNA binding show concerted motion on the microsecond-millisecond timescale. Furthermore, dynamic motion of the hinge region suggests that the DNA-binding region is capable of conformational orientations that allow it to accommodate DNA sequence variability in the cognate binding sites.**

Three major regions characterize dimeric AbrBN53: a protein dimerization region, a DNA recognition region and a looped-hinge region. The dimerization region is developed by a four-stranded intermolecular β-sheet with the dimer interface formed from strong intermolecular hydrophobic interactions. A single functional unit for DNA recognition is defined as the dimeric topology consisting of residues 13-30 (13'-30') based on mutagenic data and electrostatic character of the dimer surface. The conserved arginines 23 and 24 have been identified as being crucial for DNA recognition. Both residues are located in the central portion of the recognition region α-helices, with arginine 24 being solvent exposed and positioned for interactions with target DNA.

Arginine 8,15, 23 and 24 sidechain δ-guanido groups are localized to the central cleft and contribute to a positive charge character.

Reverse turns occurring with residues Asp 11-Glu 12, and Glu 30-Lys 31 (11'-12' and 30'- 31') are part of a linkage connecting the β-sheet scaffold and DNA-recognition region. The distance between residues Glu 12 and 12' is approximately 12 Å and can accommodate canonical double strand DNA. Electrostatic analysis of the Asp 11-Glu 12 and Glu 30-Lys 31 turns show the region to be largely electronegative . Glu 12 and Glu 30 located towards the interior of the positively-charged cleft of the DNA-recognition region, are in closest proximity to the DNA phosphate backbone. Upon binding to DNA, this region would be subject to structural displacement due to repulsive electrostatic forces. Movement of these turns away from the DNA-binding cleft would serve to minimize any unfavorable interactions resulting from the repulsive forces between the turns and the DNA phosphate backbone. If the hinge region comprised of residues Lys 9, Val 10, Lys 31 and Asp 32 of each monomeric subunit has a propensity for conformational exchange, negative steric and electrostatic effects of DNA binding would be alleviated. Such a mechanism would also contribute to an increase in the mobility for the positively-charged DNA-recognition helices.

Model for DNA-recognition. To assess more accurately the correlation between looped-hinge/DNA recognition helix mobility and protein function, ^{15}N NMR relaxation experiments were performed. The results provide one of the most compelling examples of functional mobility identified in DNA-binding proteins. We define dynamic motion as a measure of the propensity for residues to adopt multiple conformations. Identification of networks of residues with amplitudes of dynamic motion on similar timescales can define regions important for protein function. The clustered pattern of backbone dynamics identified in AbrBN53 suggests that the ability for AbrB to recognize DNA sequences with no apparent homology can be attributed to a synchronization of motion between structurally independent regions of the protein.

The area of the dimeric interface between the DNA-recognition helices is small and loosely coordinated as assessed by ps-ns motional parameters for those residues. Leu

26-Gly 27 (26'-27') are the only residues of this region exhibiting intermolecular packing interactions. Because few intermolecular contacts are present, the DNA-recognition regions of AbrBN53 are less conformationally constrained. Because few packing interactions between the DNA-recognition regions of each monomer are detected, any motion present in one monomer is independent of the other monomeric subunit. Approximately 83% of residues present in the DNA-recognition region exhibit slow and intermediate conformational exchange or rapid fluctuations on the ps-ns timescale. However, extensive intramolecular packing interactions contribute to the development of a network of residues that are capable of concerted motion. Both hydrophobic (Leu 13/13', Ile 20/20') and basic side chains (Arg 15/15', Arg 23/23', and Arg 24/24') are situated for interactions with the DNA phosphate backbone or specific nucleotide base contacts. The additional degrees of conformational freedom afforded to the DNA recognition helices from their inherent mobility can be used for optimizing sequence specific protein/DNA base-pair contacts to highly localized DNA geometry present in individual target DNA substrates.

The function of the mobile looped-hinge regions is to provide a pivot point for the DNA-recognition region. Examination of the structure of AbrBN53 shows that for each monomer the looped-hinge region is located below the negatively charged turns 1 and 3 and the DNA-recognition region. Thus, looped-hinges are optimally situated to function as a lever with the turns on one side of the pivot point (outer edge of the protein) and the DNA-recognition region on the other (towards the inner cleft). Upon interaction with a target DNA sequence, the negatively charged turns 1 and 3 (1', 3') are proposed to move out and away from the DNA phosphate backbone. This motion is made possible by the loosely constrained residues, as determined from relaxation parameters, of the hinge residues. The DNA-binding region residues located towards the interior, hydrophobic core are loosely packed and allow for the pattern of flexibility seen in the unbound protein. When binding to DNA, these residues may become less associated with the packing core, resulting in the capacity of the positively charged helices to reorient and conform to a specific DNA target sequence. The lack of extensive packing interactions between residues of the hinge region and the DNA-recognition region suggests that

motions of each region are independent but synchronous. The likelihood of the DNA-recognition region becoming structurally less restricted is supported by the observed changes in packing interactions to the β-sheet scaffold when bound to a natural promoter sequence. Thus, slight, target-dependent alterations in local conformation of the DNA-recognition region can take place to provide optimal protein side-chain/DNA base contacts.

DNA-binding proteins incorporate a remarkable variety of mechanisms for recognition and binding to target DNA sequences, often recognizing some degree of symmetry or consensus in DNA target sequences. Elucidation of the novel topology of the AbrB DNA-binding domain shed light on a general architecture of DNA recognition. However, because AbrB regulates many genes sharing no obvious target sequence consensus, it was not clear what specific mechanisms present in the protein enabled broad based, high-affinity binding. For example, proline or aromatic residues in IHF or TATA binding protein are used to facilitate binding via intercalation are not present on the side upon which DNA binding occurs. Turns within helix-turn-helix motifs or extended, solvent exposed loops displaying high B-factors or dynamic mobility linked to protein function also are not constituents of the AbrB DNA-binding domain. In this study, it is shown that independent and dynamic regions of AbrB form a global, functional network for specifically binding unrelated DNA target substrates.

Proteins that are capable of regulating a wide array of genes may help to overcome lag time in an adaptive response to environmental stress by making the response more efficient. DNA sequence discrimination and affinity, however, pose serious hurdles to be overcome for tight regulatory control. With the ability for the orientation of the looped-hinge helix DNA recognition fold to be altered, AbrB can become a specific regulatory protein for many unrelated promoter elements. Indeed, the specific structural changes occurring as a result of DNA binding can be understood only in light of what mechanistic properties inherent in the protein allow for such apparent promiscuity in DNA target recognition.

Summary

This model of AbrB structure and function outlines a general mechanism of regulatory control exerted by AbrB and recently noted homologous proteins. The key to the model is that the flexibility of the novel looped-hinge-helix DNA-binding motif allows for notable rearrangement upon exposure to different DNA targets. In this way the DNA binding properties of AbrB, and likely other transition-state regulators, can tailor itself to whatever nucleotide sequence presents itself.

To inhibit such motion and therefore impede DNA binding opens up a new approach for the design of anti-infectives. Funds to continue this line of extremely productive research are not available from this ARO award due to its untimely cancellation.

Publications:

A correlation between protein dynamics and protein:protein interactions: Millisecond motions contribute to the biological role of the response regulator Spo0F.
V.A. Feher & **J. Cavanagh**, *Nature* 400, 289 (1999).

Two-dimensional relayed-rotating-frame Overhauser spectroscopy ^{1}H NMR experiments for the selective identification of 1,2-glycosidic linkages in polysaccharides.
J.F. Cipollo, R.B. Trimble. M. Rance & **J. Cavanagh**, *Anal. Biochem.* 278, 52 (2000).

May the driving force be with you - whatever it is
J. Cavanagh & M. Akke, *Nature Structural Biology* 7, 11-13 (2000).

Novel DNA-binding domain and genetic regulation model of Bacillus subtilus transition-state regulator AbrB
J. Vaughn, V. Feher S. Naylor, M. Strauch & **J. Cavanagh**, *Nature Structural Biology* 7, 1139 (2000).

The DNA-binding domain in the Bacillus subtilis transition-state regulator employs significant motion for promiscuous DNA recognition.
J. Vaughn, V. Feher C. Bracken & **J. Cavanagh**, *J. Mol. Biol.* 305, 429 (2001).